

Nernst's heat theorem for Casimir-Lifshitz free energy

Simen A. Ellingsen

Department of Energy and Process Engineering, Norwegian University of Science and Technology, N-7491 Trondheim, Norway*

(Dated: July 28, 2008)

By regarding the Lifshitz expression for the Casimir free energy on the real frequency axis rather than the imaginary Matsubara frequencies as is customary, new light is shed on the ongoing debate regarding the thermodynamical consistency of this theory in combination with common permittivity models. It is argued that when permittivity is temperature independent over a temperature interval including zero temperature, a cavity made of causal material with continuous dispersion properties separated by vacuum cannot violate Nernst's theorem (the third law of thermodynamics). The purported violation of this theorem pertains to divergencies in the double limit in which frequency and temperature vanish simultaneously. While any model should abide by the laws of thermodynamics within its range of applicability, we emphasise that the Nernst heat theorem is a relevant criterion for choosing amongst candidate theories only when these theories are fully applicable at zero temperature and frequency.

PACS numbers: 05.30.-d, 42.50.Nn, 12.20.Ds, 65.40.Gr

Since Boström and Sernelius first predicted the existence of large thermal corrections to the Casimir force in 2000[1], controversies over the thermal behaviour of this effect, which in its most typical embodiment may be seen as the attraction between macroscopic objects due to zero-point fluctuations of the quantum vacuum, have been extensively covered in the published literature. The use of the Drude model to describe dielectric permittivity, employed in [1] was soon criticised on thermodynamical grounds [2]. The reason was that in the case of a perfect crystal lattice, when all dissipation is due to scattering of electrons on thermal phonons, the Casimir free energy as calculated with the Lifshitz formula appears to violate the Nernst's heat theorem which states that entropy should vanish as $T \rightarrow 0$. The Drude model was defended by other authors [3, 4, 5] who argued that since the Drude model offers better description of impure metals, and since real metal samples always have impurities, the Drude model must be employed. It was shown in 2003 [6] and recently in a more extensive treatment[7] that the free energy with Drude permittivity is quadratic in T for small temperatures when impurities are present. No consensus has yet been reached on the important physical question of why Casimir force predictions for the perfect lattice model, important in solid state physics, differs significantly from those pertaining to real metals with a very small but nonzero concentration of imperfections.

Recently, a somewhat analogous problem was brought forth for dielectrics with a small conductivity for finite T which vanishes at $T = 0$ [8]. While the purported violation in the case of Drude metals referred to the transverse electric (TE) mode, this time the bother appears to be a discontinuity in the transverse magnetic (TM) Fresnel reflection coefficient giving rise to nonzero entropy at zero temperature. The problem was recently argued to extend

to insulators, intrinsic and lightly doped semiconductors as well as Mott-Hubbard semiconductors and indeed the permittivity contribution from Debye rotation of molecular dipoles[9].

According to the argument presented herein both the Drude model and the dielectric permittivity model with conductivity included belong to a group of permittivities which cannot violate Nernst's theorem when permittivity can be regarded as temperature invariant in a range of temperatures near and including $T = 0$.

While the present paper does not aspire to solve the physical question of how to take into account the presence of a small conductivity in dielectrics when substituted into the Lifshitz formula, it seeks to illuminate the ever recurring question of thermodynamical consistency. It has previously been shown[10] by use of the Euler-Maclaurin (or equivalently Abel-Plana) formula that for Fresnel reflection coefficients which are continuous functions of *imaginary* frequencies in the limit $T \rightarrow 0$, Nernst's theorem is satisfied. An exploration of Casimir entropy in the formalism of surface modes was undertaken independently of this work by Intravaia and Henkel [11] whose conclusions accord with ours. By a method of summation of the eigenmodes of the vacuum between two plates they demonstrate that Nernst's theorem can be broken between metal plates only for temperature dependent relaxation such as in a perfect and infinitely large metal lattice.

This paper demonstrates a similar result using the real frequency Lifshitz formalism between plates of a generic nonmagnetic materials whose permittivity satisfies a small set of criteria. The real frequency formalism is more complicated and less elegant, but with the advantage of a more direct physical interpretation. Finally, a discussion of the physical implications of the mathematical limits involved is given. In particular we emphasise the importance of assessing when Nernst's theorem, which concerns zero temperature, can be used to inform finite temperature physics.

*Electronic address: simen.a.ellingsen@ntnu.no

I. INTRODUCTION: FREE ENERGY AND ENTROPY AT REAL FREQUENCIES

The Lifshitz expression [12] for the free energy per unit transverse area between two identical dielectric plates separated by vacuum is in general of the form

$$\mathcal{F}(a, T) = \int_0^\infty d\omega \coth\left(\frac{\omega}{2\omega_T}\right) \Im\{\phi(\omega, T)\}, \quad (1)$$

where $\omega_T \equiv k_B T/\hbar$ and $\phi(\omega, T)$ is the zero temperature integrand

$$\phi(\omega, T) = \frac{\hbar}{4\pi^2} \int_0^\infty dk_\perp k_\perp \sum_{q=\text{TE}}^{\text{TM}} \ln D_q(\omega, c\mathbf{k}_\perp, T) \quad (2)$$

wherein

$$D_q \equiv 1 - r_q^2 \exp(-2\kappa_0 a); \quad (3a)$$

$$r_{\text{TE}} = \frac{\kappa_0 - \kappa}{\kappa_0 + \kappa}; \quad r_{\text{TM}} = \frac{\epsilon\kappa_0 - \kappa}{\epsilon\kappa_0 + \kappa}. \quad (3b)$$

We have assumed $\mu = 1$ everywhere for simplicity, ϵ is the permittivity of the dielectric relative to vacuum and

$$\kappa_0 = (k_\perp^2 - \omega^2/c^2)^{1/2}; \quad \kappa = (k_\perp^2 - \epsilon\omega^2/c^2)^{1/2}. \quad (4)$$

We assume that ϵ does not depend on transverse momentum, thus neglecting any nonlocal effects and furthermore that it is a generalised susceptibility and obeys causality, which implies in particular that [13]

1. $\epsilon(-\omega^*) = \epsilon^*(\omega)$
2. $|\Im\{\epsilon(\omega)\}| > 0$ on the entire real frequency axis except at $\omega = 0$ where it may be undefined.

In general ϵ is also temperature dependent, making for the temperature dependence of $\phi(\omega, T)$. The complex conjugate is denoted with an asterisk and we will consider only real frequencies henceforth. One might furthermore impose the physically reasonable demand that

3. $\epsilon(\omega)$ is continuous and $|\epsilon(\omega)| < \infty$ for all real frequencies except possibly $\omega = 0$.

The function ϕ obeys the symmetry property $\phi(-\omega) = \phi^*(\omega)$ for real frequencies [35], hence the real part of $\phi(\omega)$ is even with respect to ω whilst its imaginary part is odd. This allows us to write \mathcal{F} in a form which makes the mathematical discussion in the following somewhat more transparent. Since both $\Im\phi$ and $\coth(\omega/2\omega_T)$ are odd functions of ω , the integrand of (1) is even and we can let the ω integral run from $-\infty$ to ∞ and divide by 2. Adding the real part of ϕ by substituting $\Im\phi \rightarrow \phi/i$ will make no difference, since it makes for an odd integrand term which vanishes under symmetrical integration, so

$$\mathcal{F}(a, T) = \frac{1}{2i} \int_{-\infty}^\infty d\omega \coth\left(\frac{\omega}{2\omega_T}\right) \phi(\omega, T) \quad (5)$$

is equivalent to (1).

Assume for the moment that ϵ , and hence ϕ , is invariant with temperature over at least a finite range of small temperatures including $T = 0$. In this case the temperature dependence of $\mathcal{F}(T)$ can be treated very simply when T is in this range, since the T dependence now sits only in the factor $\coth(\omega/2\omega_T)$.

From thermodynamics the Casimir entropy in the cavity, S , is given as

$$S = -\frac{\partial}{\partial T} \mathcal{F}(T), \quad (6)$$

so if one were able to interchange integration and differentiation with respect to T , one could write

$$\begin{aligned} S &= -\frac{1}{2i} \int_{-\infty}^\infty d\omega \phi(\omega) \frac{d}{dT} \coth\left(\frac{\omega}{2\omega_T}\right) \\ &= -\frac{\hbar}{ik_B T^2} \int_{-\infty}^\infty d\omega \phi(\omega) \frac{\omega \exp(\omega/\omega_T)}{[\exp(\omega/\omega_T) - 1]^2}. \end{aligned} \quad (7)$$

For any finite ω the integrand of (7) vanishes extremely fast, as $\exp(-\hbar|\omega|/k_B T)$, when $T \rightarrow 0$. This demonstrates the finding of Torgerson and Lamoreaux [14] that temperature corrections are important only for frequencies below ω_T , which is a very low frequency even at room temperature ($\sim 10^{13}$ rad/s). At $\omega = 0$ (and finite T) the rightmost fraction in (7) has a simple pole, yet only the imaginary part of $\phi(\omega)$ contributes to (7), which is zero here since $\Im\{\phi\}$ is an odd function of ω , removing this pole. In the limit $T \rightarrow 0$, thus, entropy vanishes as it should and the third law of thermodynamics is obeyed (there are subtleties pertaining to the TM mode as will be discussed in the following).

Two questions arise from this consideration: Under which circumstances may differentiation be interchanged with integration? And what happens if one or more parameters of $\epsilon(\omega)$ are temperature dependent all the way down to zero temperature?

II. TEMPERATURE INDEPENDENT $\epsilon(\omega)$

We will treat the first question first. Leibnitz' integral rule for improper integrals,

$$\frac{d}{dy} \int_{x_0}^\infty dx f(x, y) = \int_{x_0}^\infty dx \frac{d}{dy} f(x, y), \quad (8)$$

is always valid when ([15] §4.44)

- $f(x, y)$ and $df(x, y)/dy$ are both continuous on $x \in [x_0, \infty)$ and the relevant interval of values of y ,
- the integral on the left exists, and
- the integral on the right converges uniformly.

The generalisation to integrals with both limits infinite is trivial.

To make our considerations more concrete, let us concentrate on some permittivity models which are in common use:

$$\epsilon(\omega) = 1 - \frac{\omega_p^2}{\omega(\omega + i\nu)} \quad (9)$$

$$\epsilon(\omega) = 1 + \frac{\epsilon_\infty - 1}{1 - \omega^2/\omega_0^2 - i\gamma\omega/\omega_0^2} + \frac{i\sigma}{\epsilon_0\omega} \quad (10)$$

of which the former is the Drude model for metals, and the latter describes a semiconductor. Here ω_p is the plasma frequency, ν the relaxation frequency, ϵ_0 the vacuum permittivity and ϵ_∞, γ and ω_0 material parameters. σ is the DC conductivity of the semiconductor. Some common models which obey 1 but not 2 are

$$\epsilon(\omega) = 1 - \frac{\omega_p^2}{\omega^2} \quad (11)$$

$$\epsilon(\omega) = 1 + \frac{\epsilon_\infty - 1}{1 - \omega^2/\omega_0^2}, \quad (12)$$

the plasma model[36] for metals and a model of dielectrics with δ function dissipation at $\omega = \omega_0$ [17]. Notice that (9) and (10) obey criteria 1-3.

A. Propagating and evanescent waves

One notices that we seem to run into trouble with the continuity criterion at $|\omega| = ck_\perp$, where $\kappa_0 = 0$, since, when regarded as a double integral, (5) seems at first glance to imply integration *across* the lines $\omega = \pm ck_\perp$, which would cause trouble with continuity: one sees from (3a) that $D_q = 0$ for $\kappa_0 = 0$, hence the real part of $\text{Ln}D_q$ is undefined and the imaginary part turns out to be discontinuous as these lines are crossed.

The problem can be avoided, however. Let us define $\beta = k_\perp c/\omega$ for short. For positive frequencies, $\beta = 1$ is the limit in which the electromagnetic fields in the cavity travel parallel to the plates and become evanescent in vacuum as the $\beta = 1$ barrier is crossed, a limit whose discontinuous properties are physically obvious: the waves just on the propagating side (ck_\perp just smaller than ω) travel through the system just gracing the surfaces, while the fields on the evanescent side stay *on* the surfaces; they are qualitatively different phenomena and the transition from one to the other can be expected to be discontinuous [37]. Negative frequencies have no direct physical meaning, hence the terms “propagating” and “evanescent” must be understood in a mathematical sense here, defined by $|\beta| < 1$ and $|\beta| > 1$ respectively.

In the original Lifshitz paper[12], the \mathbf{k}_\perp integral is split automatically into propagating and evanescent parts by substituting $p = i\kappa_0 c/\omega$. Propagating contributions correspond to integrating p from 1 to 0 and evanescent to an integral from $i0$ to $i\infty$, thus avoiding the problem. We notice furthermore that the issues related to $|\beta| = 1$ occur for *any* choice of $\epsilon(\omega)$ hence can have nothing to do with the problems with Nernst’s theorem which all concern particular permittivity models.

B. Continuity

In the classical treatment by Casimir, the vacuum energy shift was found by summing over the cavity modes of the system[18], a method developed further by van Kampen et alia by use of the so-called generalised argument principle[19] and elaborated by Barash and Ginzburg[20]. The normal modes of the cavity solve the characteristic equation of the set of electromagnetic boundary conditions which reduce to the equation $D_q = 0$. At these frequencies $\varphi(\omega)$ would have poles which would cause trouble with continuity.

With permittivity models such as (9) and (10) where dissipation is included (i.e. $\epsilon(\omega)$ has a nonzero imaginary part), $D(\omega, k_\perp) = 0$ has no real-frequency solutions [38] except possibly $\omega = 0$ since $r_q \neq 1$ everywhere. The same is the case with (12) if an imaginary term is inserted in the denominator as in (10) (otherwise $|r_q| = 1$ at $\omega = \omega_0$). The real-valued permittivities as given in (11) and (12) cause r_{TM} to diverge where $\epsilon\kappa_0 + \kappa = 0$, however, in transgression of the continuity criterion.

We conclude that the continuity of $\phi(\omega)$ is ensured for all $\omega \neq 0$ so long as r_q^2 is finite, continuous and $\neq 1$, sufficient criteria for which are that $\epsilon(\omega)$ satisfies criteria 1 - 3.

What remains is the point $\omega = 0$. A priori, this is the interesting limit, since when T is very small, the coth function in (5) differs from unity only very close to zero frequency. As is well known, reflection coefficients are occasionally ill defined in the limit where ω and ck_\perp both approach zero, as is the case for the Drude model TE reflection coefficient, for example. r_q^2 is always *bounded*, however, so the integrand of $\phi(\omega)$ approaches zero in this limit due to the factor k_\perp stemming from the isotropic infinitesimal $d^2k_\perp = 2\pi k_\perp dk_\perp$ for any $\beta \neq 1$. Hence $\phi(\omega)$ is continuous for all ω if $\epsilon(\omega)$ obeys criterion 3.

A more serious problem is caused by the simple poles of $\coth(\omega/2\omega_T)$ and its T -derivative at $\omega = 0$. As argued previously, the imaginary part of $\phi(\omega)$ is zero at $\omega = 0$, so the integrand of (7) does not diverge, but is in some cases finite in this limit. For sufficiently small ω and finite σ , (9) and (10) both have the form $\epsilon \sim A + iB/\omega$ where A and B are constants, while if $\sigma = 0$ (10) instead has the form $\epsilon \sim A + iB\omega$. In both cases the imaginary part of r_{TE}^2 falls off quickly, as ω^3 and ω^5 respectively, but when $A \neq 1$, $\Im\{r_{\text{TM}}^2\}$ decreases only linearly. One easily verifies that with respect to ω , $\Im\{\phi(\omega)\} \propto \Im\{r_q^2(\omega)\}$ to leading order, hence the TM mode term of $\phi(\omega)$ is proportional to ω in the above mentioned cases.

To see how this is troublesome, consider the functions

$$H(x, t) = x \frac{e^{x/2t} + e^{-x/2t}}{e^{x/2t} - e^{-x/2t}} \quad (13)$$

$$I(x, t) = \frac{x^2 e^{x/t}}{t^2 (e^{x/t} - 1)^2}, \quad (14)$$

which are essentially (up to a constant factor) the integrands of (5) and (7) respectively when $\phi(\omega) \propto \omega$; here

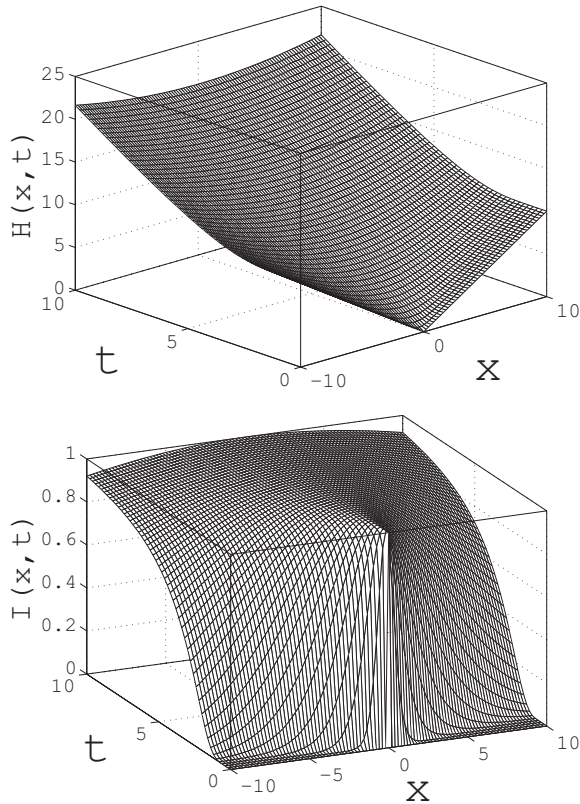


FIG. 1: Plot of (13) and (14) as functions of x and t . The functions H and I are essentially the integrands of (5) and (7) when $\phi(\omega) \propto \omega$.

x and t are suitably non-dimensionalised frequency and temperature respectively so that $\omega/\omega_T = x/t$. Notice that $\partial H/\partial t = 2I$. Equations (13) and (14) are plotted in figure 1 for x and t . If we define $H(0,0)$ to be its limiting value 0, the integrand of (5) is continuous for all ω and T as we hoped, but its T -derivative (essentially $I(x,t)$) is not. As $T \rightarrow 0$, the integrand of (7) becomes a spike of finite height and zero width. The integral past this spike is clearly zero (so the entropy would be zero as concluded above), but we run into trouble with the continuity condition. Were the limit $T \rightarrow 0$ to be taken prior to integration, $I(0,0)$ would be zero instead and I would be continuous with respect to x but discontinuous with respect to t .

Physically it does not matter whether $I(0,0)$ is 1, 0 or something in between, since the contribution from this single point is zero in either case. Hence this one point should not matter. Formally we could state this by excluding the point $\omega = 0$ from the ω integrals (5) and (7). Furthermore it is well known that the notion that *every* $\epsilon(\omega)$ on the form $A + iB\omega$ or $A + iB/\omega$ would violate Nernst's theorem is incorrect; on the contrary we argue that so long as A and B are temperature independent, none of these will. Using an analytical software such as MapleTM it is quick to check that integration of $H(x,t)$

with respect to x followed by differentiation with respect to t gives the same result as when the order of the operations is reversed. While this argument is not rigorous it should convince the reader that the continuity issues at zero temperature and frequency can be avoided since this point is of no physical significance.

A formally similar problem emerges when the permittivity is temperature dependent all the way down to zero temperature, as we will see, and in the latter case the singularity at zero temperature and frequencies *does* appear to give a physical contribution and cannot be ignored.

C. Uniform convergence

An improper integral (8) is said to converge uniformly ([15] §4.42) if $\forall \epsilon > 0$ there exists a number $a_0 > 0$ independent of T such that for all $a, a' \geq a_0$,

$$\left| \int_a^{a'} dx f(x, T) \right| < \epsilon.$$

Let us briefly analyse the behaviour of $\text{Ln}D_q(\omega, k_\perp)$ as $|\omega|$ and k_\perp approach infinity. The existence of the free energy integral (5) itself is well known, hence we need but check explicitly whether the integral (7) converges uniformly along different directions in the ω, ck_\perp plane; clearly, if the double integral over ω and k_\perp converges uniformly, the ω integral (7) does so as well.

As argued we consider propagating and evanescent contributions separately, in which case uniform convergence is straightforward to check. Reflection coefficients fall off rapidly as $|\omega| \rightarrow \infty$ (e.g. for the Drude model the real and imaginary parts of r_q^2 fall off as ω^{-4} and ω^{-5} respectively) and for $|\beta| > 1$, κ_0 is real and positive so the integrand furthermore decreases exponentially. The factor $\exp(-2\kappa_0 a)$ is oscillatory for $|\beta| < 1$, but the Dirichlet integral $\int_0^\infty dx \sin x/x$ is known to be uniformly convergent, and our integrand converges more quickly than this. It is easy to check that this also holds as $|\beta| \rightarrow 0$ and $|\beta| \rightarrow \infty$.

The splitting of (2) into propagating and evanescent parts may be done by integrating each part of the plane and taking the relevant limit to $ck_\perp \rightarrow |\omega|$ in the end. Convergence problems are then avoided for the imaginary part of $\text{Ln}D_q$ in (7); reflection coefficients fall off rapidly and further help is provided by the factor

$$\frac{\exp(\omega/\omega_T)}{(\exp(\omega/\omega_T) - 1)^2} \approx \exp(-|\omega|/\omega_T); \quad |\omega| \gg \omega_T.$$

The rate of convergence due to this factor depends on T , hence apparently cannot be used to demonstrate uniformity. We are interested only in low temperatures, however, so by defining a finite upper temperature limit \bar{T} above which the formalism is not valid, a_0 can be made T independent (dependent on \bar{T} only). The fact that the convergence of this factor alone is not uniform for infinite temperature is unproblematic, of course.

Thus we conclude that Nernst's theorem is satisfied for T independent $\epsilon(\omega)$ satisfying 1-3. The violation of Nernst's theorem in temperature dependent cases, as we shall see, can be understood as a direct consequence of violating the continuity criterion of Leibnitz' rule for improper integrals.

III. TEMPERATURE DEPENDENT PERMITTIVITY

In many models used in solid state physics, ϵ is temperature dependent for all temperatures, and herein lies the source of much of the controversy over what is the correct theory of the Casimir force between plates of real materials. The reader should note that the above theory only requires that permittivity be temperature independent for a finite temperature interval *close to zero* temperature. Rather than rigorously generalising all of the above, suffice it here to discuss how the introduction of temperature dependent permittivity illuminates the entropy problems that emerge and hints at possible resolutions. In the following we will think physically in terms of positive frequencies, bearing in mind that negative frequencies exert mathematically equivalent behaviour through the symmetry criterion 1.

The models which have caused bother so far are the TE mode of (5) using the Drude model (9) when $\nu(T) \rightarrow 0$ as $T \rightarrow 0$ (perfect lattice, no impurities) and the TM mode for dielectrics (10) with $\sigma(T) \rightarrow 0$. The two different cases share many common traits, so analysing one of them in detail will suffice as illustration. Since the Drude model has already been treated in numerous efforts by both sides of the dispute (e.g. [2, 5]), we choose the dielectric for the below discussion.

As should be clear by now, the troubles with entropy emerge for small frequencies at low temperatures. Let us from now on consider the only interesting frequency range in which $\omega^2, \gamma\omega \ll k_\perp^2 c^2, \omega_0^2$, but making no assumptions about the relative magnitude of ω and σ/ϵ_0 . Physically this corresponds to bringing ω and T close enough to the limit so that for all quantities which depend on their absolute values separately they may be replaced by zero, and only quantities which depend on their relative values, specifically the TM reflection coefficient, remains in question. In this case r_{TM} with (10) inserted simplifies to

$$r_{\text{TM}} \approx \frac{\frac{i\omega\epsilon_0}{\sigma}(\epsilon_\infty - 1) - 1}{\frac{i\omega\epsilon_0}{\sigma}(\epsilon_\infty + 1) - 1} = \frac{iv(\epsilon_\infty - 1) - 1}{iv(\epsilon_\infty + 1) - 1} \quad (15)$$

where $v \equiv \omega\epsilon_0/\sigma \equiv x/s$ where x and s are again suitably non-dimensionalised variables proportional to ω and σ respectively. We have plotted the real and imaginary part of the squared reflection coefficients as is shown in figure 2 for illustration using $\epsilon_\infty = 11.66$ as reported for Si in [22].

We find that $\Re\{r_{\text{TM}}^2\}$ is $r_0^2 \equiv (\epsilon_\infty - 1)^2/(\epsilon_\infty + 1)^2 \approx 0.71$ except for $x \ll s$ ($v \approx 0$) where it is unitary. Like-

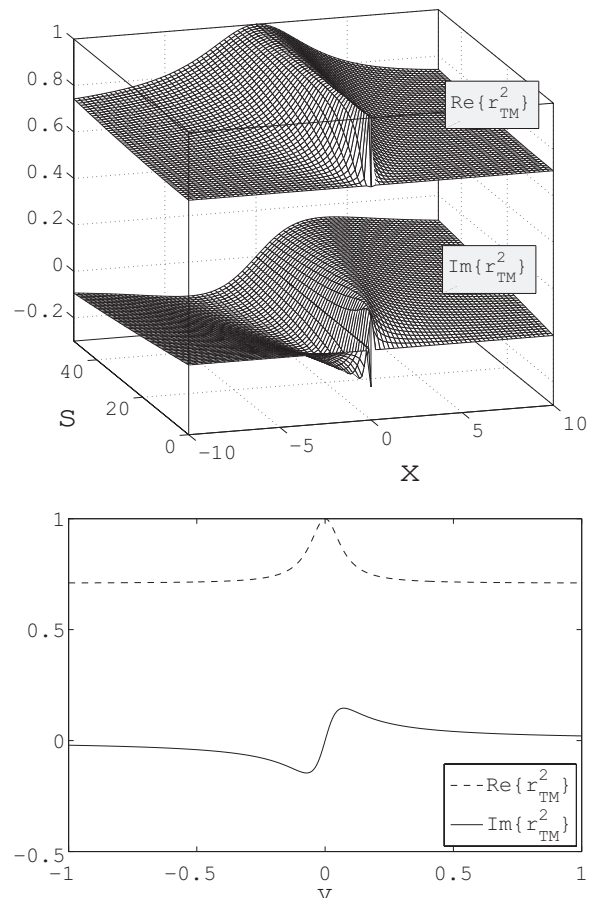


FIG. 2: Plots of the real and imaginary parts of r_{TM}^2 with (10) for $\omega, \sqrt{\omega\gamma} \ll k_\perp c, \omega_0$ and $\omega \sim \sigma/\epsilon_0$ as simplified in (15) with $\epsilon_\infty = 11.66$. Top: as function of independent x and s . Bottom as function of $v = x/s$.

wise $\Im\{r_{\text{TM}}^2\}$ for small s is approximately zero for the most part but increases to an extremum for small $|v|$ and thence decreases linearly through 0 at $v = 0$, the same linear behaviour that led to $\Im\{\phi(\omega)\} \propto \omega$ and the discontinuity of figure 1 in the TM case before, which we argued was not essential. This time, however, there are *additional* discontinuities as $s \rightarrow 0$ (equivalent to $\sigma \rightarrow 0$). In particular $\Im\{r_{\text{TM}}^2\}$ (which contributes to the integrals) is 0 everywhere except at $x = 0$ where it can take any value between its maximum and minimum ($\approx \pm 0.079$ for $\epsilon_\infty = 11.66$).

Now remember that the imaginary part of the squared reflection coefficient shown in figure 2 is to be multiplied with either the coth factor or its T derivative, both of which diverge as T/ω as $\omega \rightarrow 0$. The result is an exceedingly volatile behaviour of the integrands of (5) and (7) near zero frequency and temperature, and the limit where both are zero can take any value between $-\infty$ and ∞ depending on the way the limit is taken! This contrasts the bounded discontinuity shown in figure 1 in the temperature independent case.

Furthermore, when ϵ contains temperature dependent

parameters, entropy (7) will have an additional term

$$\frac{\hbar}{8\pi^2 i} \int_{-\infty}^{\infty} d\omega \sum_{q=\text{TE}}^{\text{TM}} \frac{e^{-2\kappa_0 a} \coth \frac{\omega}{2\omega_T}}{1 - r_q^2 e^{-2\kappa_0 a}} \cdot \frac{\partial(r_q^2)}{\partial T} \quad (16)$$

Additional entropy problems stem from this term. From (15) one finds with a little algebra that

$$\frac{\partial}{\partial T}(r_{\text{TM}}^2) = -4iv \frac{iv(\epsilon_{\infty} - 1) - 1}{[iv(\epsilon_{\infty} + 1) - 1]^3} \frac{1}{\sigma} \frac{\partial \sigma}{\partial T}. \quad (17)$$

Assuming conductivity at low temperatures to behave as $\sigma_0 \exp(-T_0/T)$ with σ_0 a constant (see below),

$$\frac{1}{\sigma} \frac{\partial \sigma}{\partial T} = \frac{T_0}{T^2}.$$

When $\omega \neq 0$ this inverse quadratic temperature dependence is no problem since (17) varies as $v^{-1} \propto \sigma$, so the term (16) is zero by a good measure when $T = 0$. The limit $\omega \rightarrow 0$ may be taken so that v has a finite value, however, in which case the derivative (17) diverges as T^{-2} . This corresponds to r_{TM}^2 making a sudden jump from 1 to r_0^2 at $T = 0$ for $\omega = 0$. The term $\partial(r_q^2)/\partial T$ at $T = 0$ is thus zero for all frequencies except $\omega = 0$, where it is infinite. We can no longer argue that this one point does not contribute to the physical quantity S , and while the purported zero temperature entropy may be difficult to calculate in this way (it is straightforward to calculate it in the imaginary frequency framework), it seems likely that entropy obtains a finite value assuming the formalism may be taken at face value.

A. Findings of the mathematical analysis

An important realisation is thus that a violation of the third law of thermodynamics is predicted in the present framework when $\epsilon(\omega)$ changes the power of its leading order term with respect to ω at *exactly* zero temperature. When $\epsilon(\omega)$ changes from $\propto \omega^{-2}$ to $\propto \omega^{-1}$, a violation occurs in the TE mode; when the change is from $\propto \omega^{-1}$ to $\propto \omega^0$, the TM mode gives the zero temperature entropy contribution.

The bottom line is that when viewed in the framework of real frequencies, all apparent zero temperature entropy anomalies stem from divergencies of the Lifshitz integrand in the double limit $T \rightarrow 0$ and $\omega \rightarrow 0$. The reader should note that in this author's understanding, this result does not contradict the findings of either [1, 3, 4, 5, 6, 7] or [2, 8, 9]. The latter of these references also notes a formal violation of Nernst's theorem due to rotation of permanent ionic dipole moments in the materials. In light of the above we may conclude that the latter effect is only problematic in media where the rotational degree of freedom of the ions vanishes with temperature in a manner so that its resonant frequency is finite at $T > 0$ and zero at $T = 0$. This anomaly needs to be studied further in the future.

B. Physical discussion of thermodynamical anomalies

In this section we will undertake a brief physical discussion of the mathematical results in the previous sections, reviewing the temperature debate for metals and semiconductors in light of the above analysis.

Models used when studying the physics of real systems are founded on assumptions which we may categorise as modelling idealisations and approximations in the description of the behaviour of these models, and there may be a need to distinguish between these in the present context. A relevant modelling idealisation, for example, is that a metal sample has infinitely large dimensions and a perfect crystal lattice structure. Much of science is founded on such ideal models and corrections to them. A relevant approximation in this context is the use of simple dielectric functions such (9) to (12) which in particular assume that the media in question have local dielectric response (i.e. they only depend on frequency, not momentum \mathbf{k}). Even for idealised systems, such approximations typically have limitations.

An ideal model which can in principle be realised (notwithstanding infeasibility in practice) cannot be allowed to violate the laws of nature, thermodynamics in particular. An approximation, on the other hand, will typically have a finite range of applicability, and cannot be expected to behave correctly outside this range. Given that the limits $T = 0$ and $\omega = 0$ are in some ways extreme cases, it is especially important to investigate the latter point in relation to the purported problems with the third law of thermodynamics.

Specifically, if an approximation which works well at room temperature does not hold for $T = 0$ one cannot conclude from a formal violation of Nernst's theorem that it cannot be used *within* its applicability range.

The much investigated temperature anomaly for metals is a good example of the above, and we will review it briefly for illustration. For a perfect and infinitely large metal lattice, relaxation $\nu(T)$ is due to electron-phonon interactions only and follows the Bloch-Grüneisen formula (see appendix D of [3]), according to which ν vanishes as T^5 as temperature tends to zero leading to the above reported anomalies. It has been pointed out that no real metal sample is ever perfect [3] nor infinitely large [39], so relaxation does not vanish in real systems. There is now consensus that for impure metals the Drude relation does not lead to thermodynamic inconsistencies.

However, the theoretical problem is thereby only halfway solved, because as pointed out [2] the fulfilment of the laws of nature cannot hinge upon the presence of imperfections: the ideal system should accord with thermodynamics as well. The solution according to the authors of [2] is to remove the relaxation from the Drude model and, more recently, introduce dissipation instead through a generalised plasma model [16], unfortunately at the expense of ignoring the manifest presence of relaxation at room temperature. Experiments seem to

confirm the predictions from such a procedure and rule out those implied by the use of the Drude model (9) (e.g. [23]) but the theory has not been universally accepted.

While the ideal crystal lattice, when treated in all detail, should certainly be found to abide by Nernst's theorem, the approximation that its dielectric response is well described by the local formalism has been questioned for temperatures approaching $T = 0$. Svetovoy and Esquivel [26] and Sernelius [27] conclude independently that at low temperatures nonlocal effects (the anomalous skin effect) dominate, and the local models are no longer reliable. Their analyses accounting for spatially dispersive effects reveal that, within the approximations made in [26, 27] Nernst's theorem is fulfilled independently of the presence of imperfections as it should. The spatial dispersion approach was criticised [28] on several accounts with reference to a treatment by Barash and Ginzburg many years ago [29] (see also [30]). The paradox remains that such a careful procedure (albeit not free of approximations) does not accord with available experimental data. Commendable efforts at a resolution include the recent exploits by Bimonte [31].

While mathematically analogous to the metal case, the temperature anomaly for semiconductors is physically different. Here the problem is not related to idealised models (the conductivity of isolators truly does vanish at zero T), but approximations only.

One can argue intuitively that the approximate model (10) probably cannot be taken at face value when conductivity is very small since it implies that even an infinitesimal conductivity should give rise to large thermal corrections in the Lifshitz formalism contrary to physical intuition. For isolators, by definition, conductivity vanishes at $T = 0$, and for many such materials conductivity even at room temperature is so small it would be expected to make for a minor perturbation only [40]. If (10) is a poor approximation at low frequencies as σ vanishes, its violating thermodynamic laws in this case may not be too worrisome.

A recent experiment [22] measured the force between a substrate of the semiconductor silicon and a gold sphere. The semiconductor was excited into a metallic state by a pulsed laser and it was concluded that while the model (10) was a good representation in the metallic state, the inclusion of the σ term when the material was in the poorly conducting state was excluded at 95% confidence. This conclusion may not be surprising in light of the above argument which indicates that a Drude type permittivity model overestimates the effect of a small conductivity in the Lifshitz formalism. If so, it is likely that the experimental result might be explained without reference to the Nernst theorem which concerns physics far removed from laboratory conditions.

Attempts at a more careful description of the effect

of a small density of free charges were recently made by Pitaevskii [32] and by Dalvit and Lamoreaux [33], based on the effects of Debye-Hückel screening from free charges in mean field theory. The resulting expressions do not fit experimental data well [34], and it is possible that a more detailed screening model is needed.

IV. CONCLUSION

While the Lifshitz formalism at real frequencies is much more complicated than the imaginary frequency equivalent normally considered, the consideration of quantities with direct physical interpretations sheds new light on the problem of non-vanishing Lifshitz entropy at zero temperature. We have argued that Nernst's heat theorem is not violated for any causal and continuous (except at $\omega = 0$) $\epsilon(\omega)$ which is independent of T near $T = 0$. This accords with the findings of Intravaia and Henkel [11] using a different approach. More generally, this holds for dielectric plates whose squared Fresnel TE and TM reflection coefficients are continuous for all ω and T and non-unitary except at $\omega = 0$. It follows from this that entropy anomalies previously reported pertain to the persistence of the permittivity's temperature variance all the way to zero temperature and are consequences of divergencies in the Lifshitz formalism in the double limit $\omega \rightarrow 0$ and $T \rightarrow 0$.

When considering physical consistency in such limits as a means to distinguish between candidate theories, particular care must be taken. We emphasise that approximations can only be judged based on their performance within their domain of applicability. Specifically, approximations which are invalid at $T = 0$ cannot be expected to be well behaved in this limit, hence cannot be rejected for causing violation Nernst's theorem which concerns zero temperature only. It is therefore important to verify carefully that approximate physical models probed by invoking Nernst's theorem are valid in this case. We finally argue that a recent experiment using an optically excited semiconductor can probably be explained without reference to the Nernst theorem by accounting for the presence of free charges more carefully than by the use of local Drude-type models.

Acknowledgements

The author thanks Professors Iver Brevik, Kimball Milton and Dr. Francesco Intravaia for discussions and useful comments. Further thanks to Professor Vladimir Mostepanenko for many useful remarks on the manuscript and for kindly allowing a preview of [9].

[1] M. Boström and B. E. Sernelius, Phys. Rev. Lett. **84**, 4757 (2000)

[2] G. L. Klimchitskaya and V. M. Mostepanenko, Phys.

- Rev. A **63**, 062108 (2001); V. B. Bezerra, G. L. Klimchitskaya and V. M. Mostepanenko, Phys. Rev. A **65**, 052113 (2002); F. Chen, G. L. Klimchitskaya, U. Mohideen, and V. M. Mostepanenko, Phys. Rev. Lett. **90**, 160404 (2003); V. B. Bezerra, G. L. Klimchitskaya, V. M. Mostepanenko, and C. Romero, Phys. Rev. A **69**, 022119 (2004)
- [3] J. S. Høye, I. Brevik, J. B. Aarseth, and K. A. Milton, Phys. Rev. E **67**, 056116 (2003)
- [4] M. Boström and B. E. Sernelius, Physica A **339**, 53 (2004)
- [5] I. Brevik, J. B. Aarseth, J. S. Høye, and K. A. Milton, Phys. Rev. E **71**, 056101 (2005); J. S. Høye, I. Brevik, J. B. Aarseth and K. A. Milton, J. Phys. A **39**, 6031 (2006); I. Brevik and J. B. Aarseth, J. Phys. A **39**, 6187 (2006); I. Brevik, S. A. Ellingsen, and K. A. Milton, New J. Phys. **8**, 236 (2006)
- [6] I. Brevik, J. B. Aarseth, J. S. Høye, and K. A. Milton, in *Quantum Field Theory under the Influence of External Conditions QFEXT03* edited by K. A. Milton (Rinton Press, Princeton, NJ, 2004) p. 54
- [7] J. S. Høye, I. Brevik, S. A. Ellingsen, and J. B. Aarseth, Phys. Rev. E **75**, 051127 (2007)
- [8] B. Geyer, G. L. Klimchitskaya, and V. M. Mostepanenko Phys. Rev. D **72**, 085009 (2005); J. Phys. A **39**, 6495 (2006); Int. J. Mod. Phys. A **21**, 5007 (2006)
- [9] G. L. Klimchitskaya and B. Geyer, J. Phys. A **41**, 164032 (2008)
- [10] e.g. K. A. Milton, J. Phys. A: Math. Gen. **37**, R209 (2004)
- [11] F. Intravaia and C. Henkel, J. Phys. A **41** 164018 (2008)
- [12] E. M. Lifshitz, Zh. Eksp. Teor. Fiz. **29** 94 (1955) [Sov. Phys. JETP **2** 73 (1956)]
- [13] L. D. Landau and E. M. Lifshitz *Statistical Physics Part 1* 3rd ed. (Oxford: Butterworth-Heinemann 1980) §123
- [14] J. R. Torgerson and S. K. Lamoreaux, Phys. Rev. E **70**, 047102 (2004)
- [15] E. T. Whittaker and G. N. Watson *A Course of Modern Analysis* 4th ed (Cambridge: Cambridge University Press 1962)
- [16] B. Geyer, G. L. Klimchitskaya, and V. M. Mostepanenko, J. Phys. A **40**, 13485 (2007)
- [17] B. E. Sernelius *Surface Modes in Physics* (Berlin: Wiley-VCH 2001) Chapter 2
- [18] H. B. G. Casimir, Proc. K. Ned. Akad. Wet. **60**, 793 (1948)
- [19] N. G. van Kampen, B. R. A. Nijboer, and K. Schram K, Phys. Lett. **26A**, 307 (1968)
- [20] Y. S. Barash and V. L. Ginzburg, Usp. Fiz. Nauk. **143** 345 (1984) [Sov. Phys. Usp. **27**, 467 (1984)]
- [21] B. E. Sernelius, Phys. Rev. B **74**, 233103 (2006)
- [22] F. Chen, G. L. Klimchitskaya, V. M. Mostepanenko, and U. Mohideen, Phys. Rev. B **76** 035338 (2007)
- [23] R. S. Decca et. al. Phys. Rev. D **75**, 077101 (2007) and references therein.
- [24] G. Burns, *Solid State Physics* International Edition (London: Academic Press, 1985) Chapter 9
- [25] F. Gebhard, *The Mott Metal-Insulator Transition* (Berlin: Springer 1997)
- [26] V. B. Svetovoy and R. Esquivel, Phys. Rev. E **72**, 036113 (2005)
- [27] B. E. Sernelius, Phys. Rev. B **71**, 235114 (2005)
- [28] G. L. Klimchitskaya and V. M. Mostepanenko, Phys. Rev. B **75**, 036101 (2007)
- [29] Yu. S. Barash and V. L. Ginzburg, Usp. Fiz. Nauk. **116** 5 (1975) [Sov. Phys.-Usp. **18** 305 (1975)]
- [30] V. M. Agranovich and V. L. Ginzburg *Spatial Dispersion in Crystal Optics and the Theory of Excitons* (London: Interscience Publishers 1966)
- [31] G. Bimonte, New J. Phys. **9** 281 (2007)
- [32] L. P. Pitaevskii, e-print arXiv:0801.0656 (2008)
- [33] D. A. R. Dalvit and S. K. Lamoreaux, e-print arXiv:0805.1676 (2008)
- [34] G. L. Klimchitskaya, V. M. Mostepanenko, and U. Mohideen, e-print arXiv:0802.2698 (2008)
- [35] This is easily argued: Because ϵ satisfies this relation by assumption, so, one finds, does D_q . The logarithm of a complex function is infinitely degenerate, and for $\ln D_q$ to give meaning we should interpret it as its principle value, $\text{Ln} D_q \equiv \ln |D_q| + i \text{Arg} D_q$, which incidentally also satisfies $\text{Ln} D_q(-\omega) = [\text{Ln} D_q(\omega)]^*$.
- [36] A generalised, causal form of the plasma model was recently proposed [16].
- [37] Due to the nonzero imaginary part of ϵ there are no similar problems for κ near $ck_\perp = |\omega| \sqrt{\text{Re}\{\epsilon(\omega)\}}$.
- [38] Sernelius has recently shown how the normal mode interpretation may still be applied [21].
- [39] The conductivity of very pure metals at low temperatures is found experimentally to be sample size dependent [24] so even assuming perfectly pure metal, ν still reaches a finite value when its Bloch-Grüneisen mean free path becomes comparable to sample dimensions.
- [40] For an introduction to different types of semiconductors, see chapter 1 of [25]